

# Warum geschichtete Stichproben?<sup>1</sup>

Von Ted Hodgson und John Borbowski

Montana State University, Bozeman, USA  
Email: [Hodgson@mathfs.math.montana.edu](mailto:Hodgson@mathfs.math.montana.edu)

übersetzt und bearbeitet von Joachim Engel, Ludwigsburg

**Zusammenfassung:** In diesem Aufsatz wird ein Lehransatz über geschichtete Zufallsstichproben präsentiert. Insbesondere haben die Autoren die Erfahrung gemacht, dass die vorgestellte Übung als hilfreiche Einführung in geschichtete Zufallsstichproben dient und Schülern hilft, ein konzeptionelles Verständnis dieses Verfahrens zu entwickeln. Dies erleichtert die Diskussion um die Charakteristika und Nutzungsmöglichkeiten der verschiedensten Stichprobenverfahren. Die Übung erreicht dieses Ziel auf intuitivem Wege und kann daher in der Sekundarstufe II wie auch in elementaren Statistikkursen an Hochschulen eingesetzt werden.

## Einführung

Aussagekräftige statistische Inferenz setzt Daten voraus, die für die Gesamtpopulation repräsentativ sind. Die Auswahl einer repräsentativen Stichprobe ist jedoch oft leichter gesagt als getan. Unterrichtserfahrungen (z.B. Rouncefield und Holmes, 1989) zeigen, daß Schüler, wenn sich selbst überlassen, dazu neigen, „Zufalls-“ Stichproben auszuwählen, die tatsächlich wenig zufällig sind. Daher ist es notwendig, auch Fragen der Stichprobenerhebung im Statistikerunterricht zu thematisieren.

Für die Auswahl repräsentativer Stichproben haben Statistiker ein breites Feld von Verfahren entwickelt. Die verbreitetsten und bekanntesten Methoden sind einfache Zufallsstichproben, geschichtete Zufallsstichproben (englisch: *stratified samples*), Klumpenstichproben und systematische Stichproben. Zur Illustration stellen wir uns vor, wir wollen eine Umfrage unter allen Schülern einer Stadt machen. Wir nehmen an, dass uns das Schulamt eine Liste aller Schüler zur Verfügung stellen kann. Eine *einfache Zufallsstichprobe* aus dieser Grundgesamtheit erhalten wir, indem wir per Losverfahren eine Stichprobe vom gewünschten Umfang wählen, die jedem Schüler die gleiche Wahrscheinlichkeit gibt, in die Stichprobe aufgenommen zu werden. Idealtypisch stellen wir uns den Losentscheid als Entnahme von Kugeln aus einer Urne vor. Bei einfachen Zufallsstich-

---

<sup>1</sup> Übersetzung aus Teaching Statistics v.20(Autumn 1998)3, S. 68-71  
*Stochastik in der Schule* 19(1999)2, S.32-41

proben kann es auf Grund von zufälliger Variabilität leicht vorkommen, dass bestimmte Teilgruppen der Gesamtpopulation unterrepräsentiert sind. In unserer Erhebung könnte es z.B. passieren, dass auf Grund der zufälligen Auswahl zu wenig Sonderschüler und zu viele Gymnasiasten (gemessen an ihrem Anteil in der Menge aller Schüler) in die Stichprobe aufgenommen werden. Eine *geschichtete Stichprobe* verhindert das, in dem der Stichprobenumfang für jede Schicht (Teilmenge, Untergruppe, englisch *Strata*) im Vorfeld festgelegt wird. Für jede Schicht wird eine eigene einfache Zufallsstichprobe erhoben. Die erhaltenen schichtenspezifischen Erhebungswerte werden dann zu einem Gesamtwert der Teststatistik verrechnet. Bei vielen Erhebungen dominieren praktische Überlegungen den Design von Stichproben. Listen der Gesamtpopulation sind oft nicht verfügbar (sondern nur Listen von Schulen bzw. Schulklassen). Ausserdem kann es sehr aufwendig sein, einzelne Schüler zu kontaktieren. Dann kann es sinnvoll sein, per einfacher Zufallsstichprobe Schulklassen auszuwählen und dann innerhalb der ausgewählten Klassen *alle* Schüler (*Klumpenstichprobe*) zu befragen. Eine *systematische Stichprobe* erhält man schliesslich, indem von einer Liste aller Schüler jeder  $k$ -te Schüler gewählt wird, z. B. jeder fünfte Schüler. Man beachte, dass eine systematische Stichprobe keinen Zufallsmechanismus zur Auswahl verlangt und daher auch keine statistische Inferenz erlaubt. Zur Diskussion der Vor- und Nachteile der einzelnen Verfahren muss auf die Literatur verwiesen werden.

Erfahrungsgemäss vergessen Schüler oft die Unterschiede zwischen diesen Verfahren wieder schnell. Ihr Abschneiden bei Prüfungen deutet darauf hin, dass sie zwar nach einem vorgegebenen Verfahren eine Stichprobe erheben können. Wenn ihnen noch weitere Informationen über die Gesamtpopulation gegeben werden, können sie oft auch ein geeignetes Stichprobenverfahren identifizieren. Jedoch ist, so unsere Erfahrung, ihr Verhalten ein weitgehendes Nachahmen. Fundamentale konzeptionelle Fragen wie z.B. „Warum wurden die verschiedenen Techniken überhaupt entwickelt?“ und „Welche Vorzüge bietet jedes einzelne Verfahren?“ werden oft mit blossem Staunen beantwortet. Schüler lernen zwar bereitwillig die Technik verschiedener Stichprobenverfahren, es fehlt ihnen aber oft ein konzeptionelles Verständnis dieser Vorgehensweisen.

### **Hintergrund und Orientierung**

Geschichtete Stichprobenverfahren bilden ein wichtiges statistisches Werkzeug, das höchst angemessen ist, wenn

(1.) die Gesamtpopulation in unterschiedliche Schichten eingeteilt werden kann und (2.) es Gründe zu der Annahme gibt, daß sich die Schichten in Hinsicht auf das zu untersuchende Merkmal unterscheiden.

Beispielsweise beschreiben Shiver und Borders (1996) ein „Experiment“, bei dem Forstverwalter die zu erwartende Holzernte von Nadelbäumen in einem Waldgebiet schätzen wollten. Da der Wald natürlicherweise in zwei Schichten eingeteilt werden kann – ein 72 Morgen grosses Gebiet mit Douglas-Tannen und ein 50 Morgen grosses Gebiet mit gemischten Nadel- und Hartholzbäumen - gab es Grund zu der Annahme, dass die Ernte in beiden Teilen unterschiedlich ausfällt. Daher sammelten die Forscher ihre Erntedaten separat für jede der beiden Untergruppen und kombinierten die Ergebnisse, um die potentielle Gesamternte zu schätzen. In ähnlicher Weise spielen geschichtete Stichprobenverfahren eine wichtige Rolle bei Erhebungen der Offiziellen Statistischen Ämter wie dem Statistischen Bundesamt oder dem amerikanischen Bureau of Labor Statistics (Scheaffer et al., 1990). Die Daten zur Berechnung des allgemeinen Preisindex sind beispielsweise von ihrer Natur her regionale Daten, und man muss bei seiner Errechnung verschiedene regionale Besonderheiten berücksichtigen wie geografische Eigenheiten, Bevölkerungsgrössen, prozentualer Bevölkerungszuwachs, industrielle Indikatoren, Niveau der Verstädterung, ethnische Besonderheiten etc.

Um in geschichtete Zufallsstichprobenverfahren einzuführen, konfrontieren wir Schüler mit einer unbekanntem Population (die in einer Papiertüte enthalten ist) und fordern sie auf, mit Hilfe geeigneter Stichprobenverfahren das arithmetische Mittel zu bestimmen. In unserem Kurs, der sich an Moore und McCabe's (1993) *Introduction to the Practice of Statistics* anlehnt, betrachten die Schüler zuerst einfache Zufallsstichproben und ihre Eigenschaften. Auf diesem Wissen aufbauend, stellen wir geschichtete Zufallsstichprobenverfahren und ihre Eigenschaften vor. Speziell erheben die Schüler Stichproben vom Umfang 4 mit Hilfe einfacher Zufallsstichproben und mittels geschichteter Zufallsstichproben und vergleichen die resultierenden Schätzungen des Populationsmittelwertes und die empirischen Stichprobenverteilungen. Die Gesamtpopulation (siehe Tabelle 1) besteht aus 40 Karten (20 rot und 20 schwarz), auf denen eine Zahl steht: der Gesamtmittelwert dieser Zahlen ist 18. Tatsächlich liesse sich diese Aktivität auch schon mit 10 Karten durchführen – eine Karte für jede Zahl – an Stelle von 40 Karten. Jedoch könnten beim Verwenden von lediglich 10 Karten die Schüler berechtigterweise fragen, warum hier nicht die gesamte Population als Stichpro-

be gewählt wird. Weiterhin wird die Stichprobe ohne Zurücklegen gewählt. Eine kleinere Gesamtpopulation macht es dann weniger wahrscheinlich als eine größere Grundgesamtheit, extreme Ergebnisse ('alle Karten rot') zu erhalten. Eine Grundgesamtheit vom Umfang 40 verspricht genügend unterschiedliche Stichproben und ermutigt die Verwendung anderer Stichprobenverfahren als einfache Zufallsstichproben und ist dennoch klein genug, um von den Schülern bewältigt werden zu können.

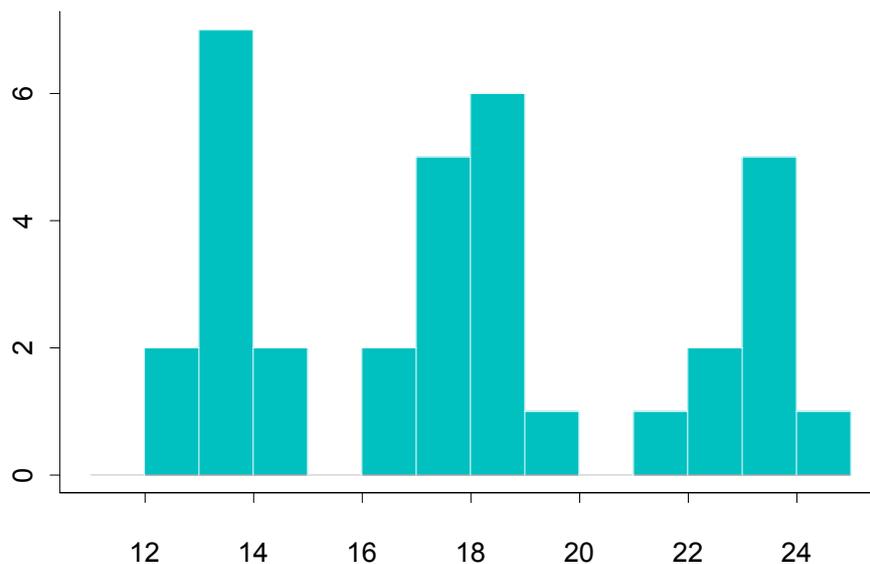
<b>X</b>	<b>HÄUFIGKEIT</b>	<b>KARTENFARBE</b>
6	4	Rot
7	4	Rot
8	4	Rot
9	4	Rot
10	4	Rot
26	4	Schwarz
27	4	Schwarz
28	4	Schwarz
29	4	Schwarz
30	4	Schwarz

**Tabelle 1:** Grundgesamtheit

Zu Beginn der Aktivität ziehen Schüler Zufallsstichproben vom Umfang 4 und errechnen das Stichprobenmittel. Die Stichproben, die kürzlich in einer Klasse mit 34 Schülern gezogen wurde, sind mitsamt den entsprechenden Stichprobenmittelwerten in Tabelle 2 wiedergegeben (geordnet nach dem Mittelwert). In der Klasse selbst sind diese Mittelwerte schnell errechnet und in Form von Tabellen und Grafiken dargestellt. Sie können daher als Basis für eine Klassendiskussion dienen. Obwohl wir die Daten 'zu Fuß' errechneten, lädt diese Aktivität auch zum Einsatz verfügbarer Technologien ein wie z.B. Tabellenkalkulationssystemen, Statistikpaketen oder grafikfähiger Taschenrechner. Im Hinblick auf unsere Schulklasse offenbarte eine anfängliche Inspektion der Ergebnisse wenig Trends. Das Stichprobenmittel erstreckte sich zwischen 12,25 und 24 und erschien (wenigstens gemäss der Tabelle) kaum um einen speziellen Wert zentriert. Zusätzliche Informationen über die zugrundeliegende Gesamtpopulation kann oft noch aus der Stichprobenverteilung entnommen werden. Die Schüler unserer Klasse kamen jedoch zu der Schlussfolgerung, dass die Verteilung keine weiteren Informationen in sich birgt. Der allgemeine Klassenkonsens war, dass hier einfache Zufallsstichproben nur wenig definitive Information über das Populationsmittel in sich tragen.

Stichprobe Nr.	Daten				Stichprobenmittel
1	26	7	10	6	12,25
2	10	26	9	6	12,75
3	29	6	7	10	13
4	6	8	9	29	13
5	5	9	8	30	13,25
6	9	8	7	29	13,25
7	7	7	30	9	13,25
8	9	9	10	26	13,5
9	9	8	8	30	13,75
10	9	10	8	29	14
11	10	9	29	9	14,25
12	6	27	6	26	16,25
13	7	7	26	27	16,75
14	28	8	6	26	17
15	7	6	29	26	17
16	6	29	26	8	17,25
17	9	6	26	29	17,5
18	26	9	8	28	17,75
19	7	10	26	29	18
20	27	6	30	9	18
21	6	29	28	10	18,25
22	8	29	26	10	18,25
23	6	8	30	30	18,5
24	26	9	30	10	18,75
25	8	11	28	30	19,25
26	26	27	6	27	21,5
27	30	26	27	6	22,25
28	8	26	29	28	22,75
29	10	26	26	30	23
30	29	6	30	27	23
31	28	9	30	26	23,25
32	27	26	30	10	23,25
33	30	10	29	26	23,75
34	29	30	7	30	24

**Tabelle 2:** Einfache Zufallsstichproben (von Schülern erhoben)



**Abbildung 1:** Ein Häufigkeitshistogramm der Mittelwerte der einfachen Zufallsstichproben

Anschliessend informieren wir die Schüler, dass das Kartendeck aus zwei Sorten von Karten besteht: rote und schwarze Karten. Auch wenn ihnen nicht mitgeteilt wird, dass sich die beiden Teilpopulationen der roten und schwarzen Karten im Hinblick auf das interessierende Merkmal (hier der numerische Wert auf der Karte) unterscheiden, wird den Schülern gesagt, dass möglicherweise ein Unterschied bestehen *könne* und dass geeignete Verfahren zur Stichprobenentnahme diese potentiellen Unterschiede berücksichtigen könnten.

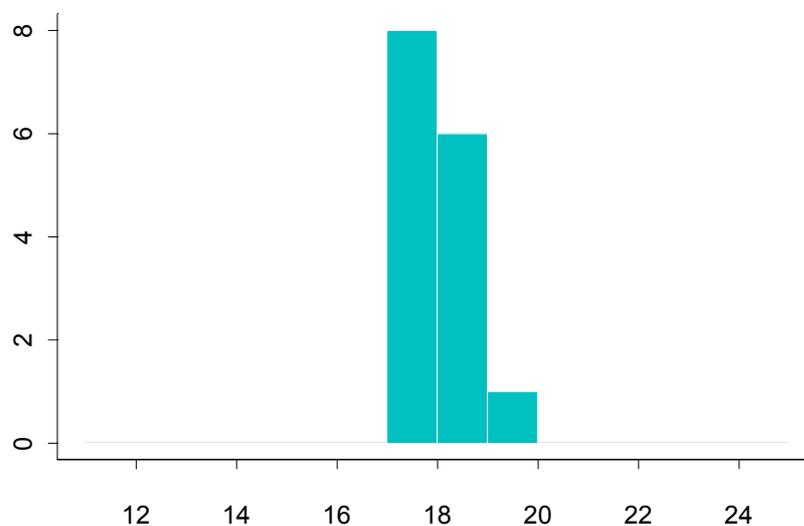
Ein solches Verfahren besteht in der Auswahl roter und schwarzer Karten, so dass der Prozentsatz von beiden Gruppen in der Stichprobe die Verhältnisse in der Grundgesamtheit widerspiegeln. Hier werden also geschichtete Zufallsstichprobenverfahren eingeführt. Da die Grundgesamtheit aus 20 roten und 20 schwarzen Karten besteht, ziehen die Schüler Stichproben vom Umfang 2 aus jeder Teilpopulation (die wir in unterschiedliche Tüten gelegt haben) und mitteln die Resultate.

Man sieht hier, dass sich die Teilpopulationen hinsichtlich des interessierenden Merkmals unterscheiden. Das heisst, roten Karten entsprechen kleine Zahlen, wohingegen der Wert der schwarzen Zahlen immer gross ist. Daher verspricht, wie die von denselben Schülern erzeugten und in Tabelle 3 dargestellten Daten zeigen, das geschichtete Ziehen von Zufallsstichproben präzisere Schätzungen des Populationsmittelwertes als einfache Zufallsstichproben.

Stichprobe Nr.	Daten				Stichprobenmittel
1	6	7	27	28	17
2	8	9	26	27	17,5
3	8	6	28	28	17,5
4	7	8	29	26	17,5
5	9	9	26	26	17,5
6	6	9	29	27	17,75
7	8	10	26	27	17,75
8	10	6	27	28	17,75
9	9	9	28	26	18
10	6	8	28	30	18
11	10	8	27	28	18,25
12	10	7	28	29	18,5
13	9	9	27	29	18,5
14	8	9	29	29	18,75
15	9	10	28	29	19

**Tabelle 3:** Geschichtete Zufallsstichproben der Schüler

In dieser Klasse wurden 15 geschichtete Zufallsstichproben gezogen und die erhaltenen Mittelwerte waren konsistent „nahe“ am Populationsmittel von 18 (das erst dann den Schülern offenbart wurde, nachdem sie geschichtete Stichproben gezogen und ihre Mittelwerte errechnet haben). Darüberhinaus ist, zumindest in diesem Fall, die Verteilung der Stichprobenmittel weitaus weniger variabel als bei einfachen Zufallsstichproben (siehe Abbildung 2). Der didaktische Effekt dieser Simulation ist daher, dass Schüler den potentiellen Zuwachs an Präzision, der mit geschichteten Zufallsstichproben einher geht, beobachten können.



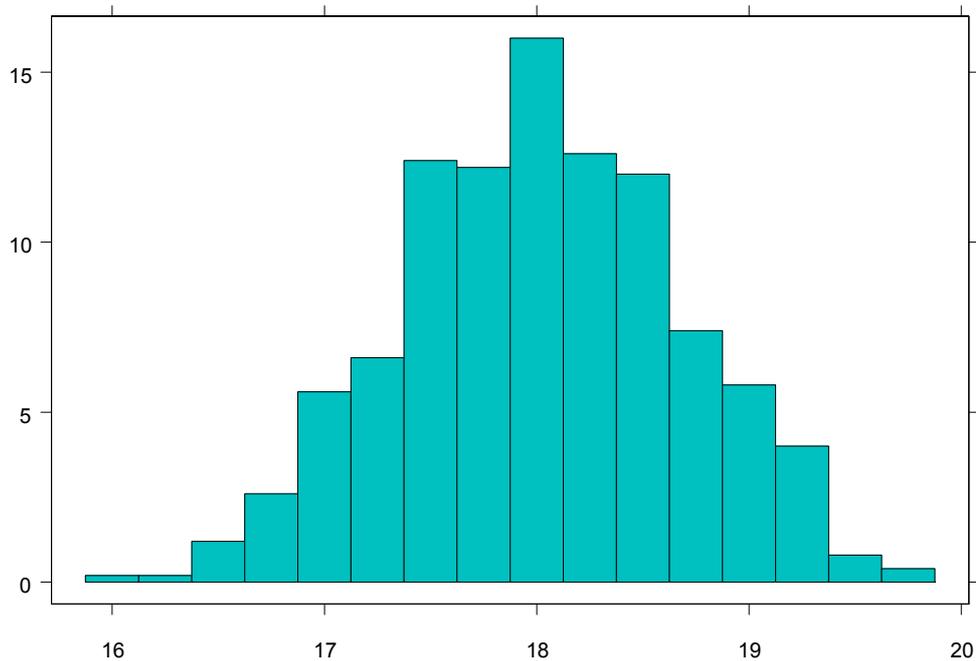
**Abbildung 2:** Ein Häufigkeitshistogramm der Mittelwerte der geschichteten Zufallsstichproben

### **Schlussfolgerungen und Ergänzungen**

Wir ziehen die Schlussfolgerung, dass Übungen mit einfachen Zufallsstichproben Schülern helfen, eine höhere Wertschätzung und ein besseres konzeptionelles Verständnis für weithin angewandte Stichprobenverfahren zu entwickeln. Die durchgeführte Übung ermutigt Schüler insbesondere, zwei kritische Charakteristika von Stichprobenverfahren zu beachten: die Präzision einer Stichprobenstatistik und die Variabilität der resultierenden Stichprobenverteilung. Darüber hinaus erwies sich der gewählte Ansatz über Simulationen als ausserordentlich reichhaltige Quelle für weitere Problemstellungen und Untersuchungen.

Im betrachteten Beispiel konzentrieren sich die vorangegangenen Überlegungen auf eine Gesamtpopulation, bei der die Anzahl roter und schwarzer Karten gleich ist. In den allermeisten realen Situationen ist jedoch die Anzahl der Individuen in allen identifizierbaren Teilpopulationen unterschiedlich gross. Daher besteht eine natürliche Erweiterung der anfänglichen Übung darin, den Prozentsatz der roten und schwarzen Kugeln zu variieren. Analog besteht eine andere Erweiterung in der Frage, was geschieht, wenn die Grundgesamtheit zwar in Schichten aufgeteilt werden kann, aber die Teilpopulationen sich tatsächlich ähnlich sind hinsichtlich des interessierenden Merkmals. Eine einfache Simulation zeigt, dass geschichtete Zufallsstichproben und einfache Zufallsstichproben in dieser Situation ähnliche Resultate hervorbringen.

Zuletzt lassen sich die Resultate der „Hand-“ Simulationen vergleichen mit Computer oder Taschenrechner- Simulationen. Beispielsweise zeigt Abbildung 3 die Resultate einer Computersimulation von 500 geschichteten Stichproben. Man beachte, dass die grössere Zahl der Stichproben eine feinere Analyse erlaubt und die wahre Stichprobenverteilung präziser offenbart. Im allgemeinen können grosse Simulationen die „wahren“ Unterschiede zwischen verschiedenen Stichprobenverfahren offenbaren. Zusätzlich können grosse Simulationen zu Diskussionen über andere statistisch wichtige Konzepte führen, wie z.B. die Auswirkung des Stichprobenumfangs auf das Resultat und das asymptotische Verhalten der empirischen Stichprobenverteilungen.



**Abbildung 3:** Ein Häufigkeitshistogramm für die Mittelwerte von 500 geschichteten Stichproben

## Literatur

- Moore, D.S. und McCabe, G.P. (1993). *Introduction to the Practice of Statistics* (Second Edition). New York: W.H. Freeman.
- Rouncefield, M. und Holmes, P. (1989). *Practical Statistics*. London: Macmillan.
- Scheaffer, R.L., Mendenhall, W. und Ott, L. (1990). *Elementary Survey Sampling*. Boston: PWS -Kent.
- Shiver, B.D. und Borders, B.E. (1996). *Sampling techniques for Forest Resource Inventory*. New York: John Wiley & Sons.

## Ergänzende Hinweise zur Literatur:

Das Buch von Scheaffer, Mendenhall und Ott ist als elementare Einführung sehr zu empfehlen, weil es klar, einfach und illustrativ geschrieben ist. Ein Klassiker zur Stichprobentheorie ist auch heute noch

Cochran, W. G. (1977): *Sampling Techniques*. 3. Auflage New York: Wiley.

*Deutschsprachige Literatur zur Stichprobentheorie:*

Kreienbrock, L. (1993): *Einführung in die Stichprobenverfahren. Lehr – und Übungsbuch der angewandten Statistik. Mit Übungsaufgaben und Prüfungsfragen*. München: Oldenbourg-Verlag

Leiner, B. (1985): *Stichprobentheorie: Grundlagen, Theorie und Technik*. München; Wien: Oldenbourg. (1994 erschien die 3., durchges. Auflage)

Schwarz, H. (1975): *Stichprobenverfahren. Ein Leitfaden zur Anwendung statistischer Schätzverfahren*. München: Oldenbourg-Verlag.

Stenger, H. (1986): *Stichproben*. Heidelberg- Wien: Physica-Verlag.